Canadian Institute of Actuaries / Institut canadien des actuaires

Research Paper

The Mathematical Mechanism of Biological Aging

Research Council –
Academic Research Committee

October 2019

Document 219106

Ce document est disponible en français

The Mathematical Mechanism of Biological Aging

Boquan Cheng, Bruce Jones, Xiaoming Liu, and Jiandong Ren
Department of Statistical and Actuarial Sciences
University of Western Ontario
London, ON, N6A 5B7 Canada

May 21, 2019

Despite aging being a universal and ever-present biological phenomenon, describing this aging mechanism in accurate mathematical terms – in particular, how to model the aging pattern and quantify the aging rate – has been an unsolved challenge for centuries. In this paper, we propose a Coxian-type Markovian model that can provide a quantitative description of the well-known aging characteristics – the genetically determined, progressive, and essentially irreversible process. Our model has a unique structure, including a constant transition rate for the aging process, and a functional form for the relationship between aging and death with a shape parameter to capture the biologically deteriorating effect due to aging. The force of moving from one state to another in the Markovian process indicates the intrinsic biological aging force. The associated increasing exiting rate captures the external force of stress due to mortality risk on a living organism.

The idea of the paper is developed from Lin and Liu (2007). A big difference is that, in this paper, our model uses a functional form for model parameters, which allows a parsimonious yet flexible representation for various aging patterns. Our proposed mathematical framework can be used to classify the aging pattern, and the key parameters of the model can be used to measure and compare how human aging evolves over time and across populations.

1. Introduction

> Human life is a very personal affair. It is your life and mine and that of our neighbor. Each life is a separate and distinct entity; yet there is a common stamp upon all ... Many lives are cut short ... an accident or an acute illness may snuff out the light before it is burned out. But burn out it will, ultimately, even if no such accidents intervene.
>
> Dublin et al. (1949)

In the above sentence, Dublin et al. have well described the fact that while individual lifetimes could have been affected by many personal and random factors, there is this ubiquitous aging process underneath the life of all human beings.

Then, what is "aging"?

To begin with, let us quote the definition of "aging" from Jones (1956):

> Aging, as applied to living organisms, is the genetically determined, progressive, and essentially irreversible diminution with the passage of time of the ability of an organism or of one of its parts to adapt to its environment, manifested as diminution of its capacity to withstand the stresses to which it is subjected (i.e. the increase of susceptibility to certain diseases with age), and culminating in the death of the organism.

Jones' definition has provided some characteristics about the aging process, such as *genetically determined*, *progressive* and *essentially irreversible*. But it was unclear on what capacities and how they decline with the passage of time in this aging definition. Like a great deal of other research at that time (before the middle of the 20th century), the studies on the aging process were mainly descriptive and not accessible for measurement due to the lack of directly observable aging-related data. For this reason, the earlier theories of aging tried to associate the impact of aging with the increasing death rates as a way to justify their hypotheses quantitatively. In other words, it was a common practice that the exponentially increasing mortality pattern (the Gompertz law of mortality, due to its good fit to a large portion of adult mortality rates) acted as a manifest of the "indirectly" observed aging process.

However, we would like to clarify the relationship between aging and death: while the mortality rate is strongly correlated to the rate of aging, the two are not exactly equivalent. There are many external factors that can cause pre-mature death, and/or even alter the aging process. To name a few: family background, education, income, lifestyle, marital status, accidents, and acute diseases. According to Herskind et al. (1995), an additive genetic component explains about 2 per cent of variability in life span, indicating that non-genetic factors make substantial contributions to life span. Rather than thinking of aging and death in terms of the cause-and-effect relationship, it is now believed that the age pattern of mortality risk results from interaction between the process of individual aging and external stresses, as stated in Yashin et al. (2012) and many references therein.

Another big change in the past few decades regarding the studies of aging is that many researchers began to collect longitudinal data of aging-related factors. These factors include body mass index, diastolic blood pressure, pulse pressure, pulse rate, level of blood glucose, hematocrit, and serum cholesterol; again see Yashin et al. (2012) and references therein for the recent research in this direction. Different from many previous studies, the observations in these studies are obtained from young healthy individuals in order to understand how the biological processes develop in aging humans even before the onset of chronic diseases. In Yashin et al. (2012), it is found that different factors display different average age patterns and different dynamic properties. Some factors follow monotonic changes with increasing age; some are non-monotonic (e.g., increase then decline). Their findings also reveal that, when an individual's physiological indices deviate from their "optimal" values, one will experience

higher-than-average mortality risk. From our perspective, their research might be more useful in understanding why some individuals can live longer life than others. However, if we want to develop a holistic quantitative tool to describe the aging effect of population, we might need to work on more general statistics.

By *more general statistics* we mean concepts similar to allostatic load defined in Crimmins et al. (2003) or the biomarker as defined in Belsky et al. (2015). In summary, these are compiled indices based on various physiological functions across multiple organ systems (e.g., pulmonary, periodontal, cardiovascular, renal, hepatic, and immune function) to measure the cumulative biological burden developed over time or to mark the deterioration degree of an aging body. Despite the intensified studies in the past few decades and better longitudinal data accessibility now, the research in this area is still in its infant stage. This is largely due to the complexity of the living and dying process. From our point of view, it seems impossible to reach a commonly accepted biological aging index in the near future.

As a result, instead of taking an absolute approach in which one needs to select one or multiple representative aging indices, we propose a Coxian-type Markovian model that can provide a quantitative description of the well-known aging characteristics – "the genetically determined, progressive and essentially irreversible" process. Our model has a unique structure, including a constant transition rate for the aging process, and a functional form for the relationship of aging and death with a shape parameter to capture the biologically deteriorating effect due to aging. The proposed model incorporates two separate yet related forces in this Markovian framework: the force of aging (i.e., the intrinsic biological process) and the force of dying (i.e., the external forces from environmental stress including accidents, accessibility of nutrition and medical care, etc.).

The idea of the paper is developed from Lin and Liu (2007). A big difference is that in this paper our model uses a functional form for model parameters, which allows a parsimonious yet flexible representation for various aging patterns.

The proposed model belongs to a big class of distributions: the so-called phase-type distribution. There is a long history of using phase-type distributions for survival modelling in the category of "absorbing time" distributions; see Aalen (1995), Asmussen et al. (1996), Lin and Liu (2007), and Su and Sherris (2012), to list a few. Phase-type distributions can be clearly implemented when the states of the underlying process are observable and can be pre-specified. Otherwise, the flexibility of phase-type distributions might turn into a disadvantage since it will make model estimation much more difficult due to the non-uniqueness property of the phase-type distribution (Slud and Suntornchost, 2014). Our defined aging process is a latent process; however, a carefully designed functional form for model parameters helps mitigate the non-identifiability problem.

Our proposed mathematical framework can be used to specify the aging pattern, and the key parameters of the model can be used to measure the deterioration effect of aging. Our model can also provide biologically meaningful comparisons regarding how human aging evolves over time and across populations. Other advantages of the model include:

- The model has a small number of parameters that can be conveniently estimated using

maximum likelihood estimation. The parsimonious structure of the model ensures that it can be uniquely identified with a given data set and the resulting model can provide a meaningful biological interpretation.

- The model is flexible in terms of how it represents the impact of aging on mortality.
- The model introduces randomness associated with aging and allows for quantifiable heterogeneity in terms of the physiological age of individuals in a population. (Note: Physiological age and biological age are interchangeable terms used to capture the extent to which an individual has aged biologically. Normally, physiological age is calibrated so that an individual who is chronological age x has a physiological age that is greater than (less than) x if the individual has aged more than (less than) an average person of chronological age x.)
- This heterogeneity allows one to capture mortality selection at older ages; that is, the tendency for less healthy lives of a given age to experience higher mortality, leaving a healthier group of survivors.

The paper is organized as follows. In Section 2, we introduce our phase-type aging model. We begin with some background on phase-type distributions followed by the characteristics of a subclass known as Coxian phase-type distributions. We then introduce our model, which is a special case of the latter, and describe the structure of the model parameters. In Section 3, we discuss how the model can be calibrated using lifetime data, and illustrate these ideas using a small data set providing lifetimes of residents of a retirement community. We further demonstrate how our model can provide useful interpretation for the embedded aging process. In Section 4, we present some further analysis of our model. We illustrate how it can capture the lifetime distribution of a very different aging mechanism. Through a simulation study, we show the challenge in estimating the number of model states with only lifetime data, and demonstrate how the number of states influences the variability in how individuals age according to the model. Section 5 concludes the paper with a summary and some suggestions for further research.

2. The proposed phase-type aging model (PTAM)

2.1. Basic mathematics for phase-type distributions

In this section, we provide the definition of phase-type distribution and describe its basic properties. We have also tried to organize the material to facilitate discussion about the challenges in using this type of model in practice. For the general theory of phase-type distributions, see, for example, Neuts (1982) and Asmussen (1989).

Definition: *Let Y_t be a time-homogeneous Markov process defined on a finite state-space $S = E \cup \Delta = \{1,2,\cdots,m\} \cup \Delta$, where Δ is absorbing and the states in E are transient. Let Y_t have initial distribution $(\boldsymbol{\alpha}^T, 0)$ (written as a row vector), and infinitesimal generator*

$$\begin{pmatrix} \mathbf{\Lambda} & \mathbf{q} \\ \mathbf{0} & 0 \end{pmatrix} \tag{1}$$

where $\mathbf{q} = -\mathbf{\Lambda e}$ and $\mathbf{e}$ is the column vector of ones.

Let T denote the time until absorption or the time until death in the human lifetime context. Then T is said to follow a phase-type (PH) distribution with representation $(\boldsymbol{\alpha}, \boldsymbol{\Lambda})$.

One very intriguing aspect of phase-type models is the intuitive interpretation that they provide about the underlying biological/engineering mechanism of the system that results in the lifetime survival distribution. In plain language, the survival time of such a system is the sum of the sojourn times in all states that the process has ever visited before it is absorbed. The survival probability at any time t is the total probability that the process is in any state other than the absorbing state at time t.

For $s \geq 0, t \geq 0, i, j \in E$, let $P_{ij}(s, s+t)$ be the transition probability from state i to state j during the time period $(s, s+t]$:

$$P_{ij}(s, s+t) = \Pr(Y_{s+t} = j | Y_s = i)\,. \tag{2}$$

Because Y_t is a time-homogeneous Markov process, the probability $P_{ij}(s, s+t)$ does not depend on s, so we write $P_{ij}(t) = P_{ij}(s, s+t)$ for all $s \geq 0$.

Denote $\mathbf{P}(t) = \{P_{ij}(t)\}_{i,j \in E}$ as the transition probability matrix (among the transient states) of Y_t in the time interval $[0, t]$. Then it satisfies the Kolmogorov forward equation,

$$\frac{d}{dt}\mathbf{P}(t) = \mathbf{P}(t)\boldsymbol{\Lambda}\,, \tag{3}$$

with the initial condition $\mathbf{P}(0) = \mathbf{I}$, where $\mathbf{I}$ is the identity matrix.

The Kolmogorov forward equation has the unique solution

$$\mathbf{P}(t) = \exp(\boldsymbol{\Lambda} t), \tag{4}$$

where the matrix-exponential $\exp(\boldsymbol{\Lambda} t)$ is defined by

$$\exp(\boldsymbol{\Lambda} t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \boldsymbol{\Lambda}^n.$$

As a result, the survival function of the time until absorption random variable T can be expressed as

$$S(t) = \boldsymbol{\alpha}^T \exp(\boldsymbol{\Lambda} t)\mathbf{e},\ t > 0\,. \tag{5}$$

Notice that the ith element of the vector

$$\mathbf{p}(t) = \boldsymbol{\alpha}^T \exp(\boldsymbol{\Lambda} t) \tag{6}$$

gives the probability that $Y_t = i$, and $\mathbf{p}(t)\mathbf{e}$ gives the probability that $Y_t \in E$.

Phase-type distributions have simple matrix expressions for their survival function, probability density function, Laplace transform, and raw moments. Furthermore, the class of phase-type distributions is dense in the space of all continuous positive distributions. That is, any distribution on $(0, \infty)$ can, at least in principle, be approximated closely by a phase-type distribution. Therefore, it has been successfully applied in many areas to obtain explicit analytical solutions. However, obtaining numerical results by solving differential equation (4) or calculating matrix exponential (5) is not trivial, especially when the dimension of the matrix $\boldsymbol{\Lambda}$ (i.e., the number of states, m) is big.

In addition, a general phase-type distribution has many parameters, and sometimes the way of parameterizing a model is not unique (Asmussen et al., 1996). Therefore, the parameter estimation of phase-type distributions is known to be difficult. As discussed in the later part of this paper, the problem of non-uniqueness in parameter specification is usually intermingled with estimation difficulties of the model, especially when the model contains a large number of states and the model parameters are close to each other.

We remark that this paper considers the application of phase-type modelling to the aging process, similar to Aalen (1995), Lin and Liu (2007), and Su and Sherris (2012). In such applications, the model may involve a large number of states but the intensity matrix is sparse and follows certain structures.

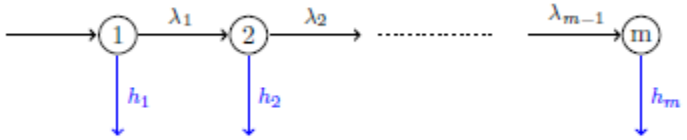
2.2. Coxian phase-type distributions

In the following, we discuss the Coxian-type distribution, which is a special type of phase-type distribution with initial distribution $\boldsymbol{\alpha}^T = (1,0,\cdots,0)$ and intensity matrix

$$\boldsymbol{\Lambda} = \begin{bmatrix} -(\lambda_1 + h_1) & \lambda_1 & & & \\ & -(\lambda_2 + h_2) & \lambda_2 & & \\ & & \ddots & & \\ & & & -(\lambda_{i-1} + h_{m-1}) & \lambda_{i-1} \\ & & & & -h_m \end{bmatrix}, \tag{7}$$

where $\lambda_i > 0$ and $h_i > 0$ for all $i \in E$. The Coxian model is illustrated in Figure 1. Since the process only starts from state 1, the survival function of the Coxian-type model only depends on the first row of the transition matrix $\mathbf{P}(t)$, which is denoted by $\mathbf{p}_1(t)$. Let the kth element of $\mathbf{p}_1(t)$ be denoted as $P_{1k}(t)$ or simply $P_k(t)$.

Figure 1: Diagram for a Coxian-type Markovian process.



Then the Kolmogorov forward equation is given by

$$\begin{cases} \frac{dP_1(t)}{dt} & = -(\lambda_1 + h_1)P_1(t); \\ \frac{dP_k(t)}{dt} & = \lambda_{k-1}P_{k-1}(t) - (\lambda_k + h_k)P_k(t),\ k = 2,3\ldots,m. \end{cases} \tag{8}$$

Obviously $P_1(t) = e^{-(\lambda_1+h_1)t}$. In addition, $P_k(t)$ for $k = 2,\ldots,m$ can be obtained iteratively, yielding

$$P_k(t) = \sum_{j=1}^{k} \frac{(-1)^{k-1}\lambda_1\ldots\lambda_{k-1}}{\prod_{s=1,s\neq j}^{k}(\lambda_j+h_j-\lambda_s-h_s)} e^{-(\lambda_j+h_j)t},\ k = 2,\ldots,m. \tag{9}$$

Finally, the survival function of T is given by

$$S(t) = \sum_{k=1}^{m} P_k(t). \tag{10}$$

When the dimension of the Coxian model is small and the parameter values are far apart so that the denominator of (9) is not too small, equation (9) may be used to quickly evaluate the distribution. However, in lifetime modelling, the model to be used may contain a large number of states with λ_is and h_is being very close to each other. Using (9) causes numerical problems because the values of each term in summation are large and have alternative signs.

2.3. A class of Coxian phase-type distributions – our proposed PTAM

The Coxian distribution provides a natural and appealing approach to modelling aging. Unfortunately, the general Coxian distribution has many parameters, making estimation difficult. We can greatly simplify the model while retaining considerable flexibility by imposing some structure. That is, we can specify mathematical forms for the pattern of λ_i values, representing the internal force of aging, and the pattern of h_i values, representing the external force of dying. If the mathematical functions used for this involve a small number of parameters, then our estimation task becomes manageable.

First, assume that the number of states (physiological ages) m in our model is known. We will address how to determine m later in the paper.

Now assume that $\lambda_i = \lambda$, a constant, for all $i = 1,2,\ldots,m-1$. In doing so, we are assuming that the aging rate is uniform over time, so that the rate of increase in physiological age is uniform. This makes some sense because calendar age increases uniformly. Although our assumed aging rate is constant, there is variability in transition times. Therefore, at any calendar age, an individual may be in a number of different states representing different physiological ages.

The value of λ needs to be appropriate to the value of m. We need the rate of progression through the states to be large enough that some individuals will reach state m, or there is no need to have as many as m states. However, we want only a small proportion of individuals to survive to state m, or the model death rate will flatten out at older ages. We therefore let

$$\lambda = m/\psi, \tag{11}$$

where ψ can be thought of as the life span of individuals in the population of interest. The parameter ψ need not be a limiting age, as there is no limiting age in our model. However, ψ

should be a high age to which only a very small proportion of individuals will survive. This parameter can be estimated from data or chosen based on prior opinion. For human lifetimes, a value of ψ between 100 and 120 may be reasonable.

We next specify a form for the h_i values. First, suppose that h_1 and h_m are fixed. These are parameters to be estimated. We then require a smooth pattern of h_i values for i between 1 and m. We achieve this by letting

$$h_i^s = \frac{m-i}{m-1} h_1^s + \frac{i-1}{m-1} h_m^s.$$

In other words, powers of h_i are obtained as a linear interpolation between the corresponding powers of h_1 and h_m. We then have

$$h_i = \left(\frac{m-i}{m-1} h_1^s + \frac{i-1}{m-1} h_m^s \right)^{1/s}.$$

We can use this for any real s except $s = 0$, when the expression is undefined. However, the limiting case as $s \to 0$ gives

$$h_i = h_1^{\frac{m-i}{m-1}} h_m^{\frac{i-1}{m-1}}.$$

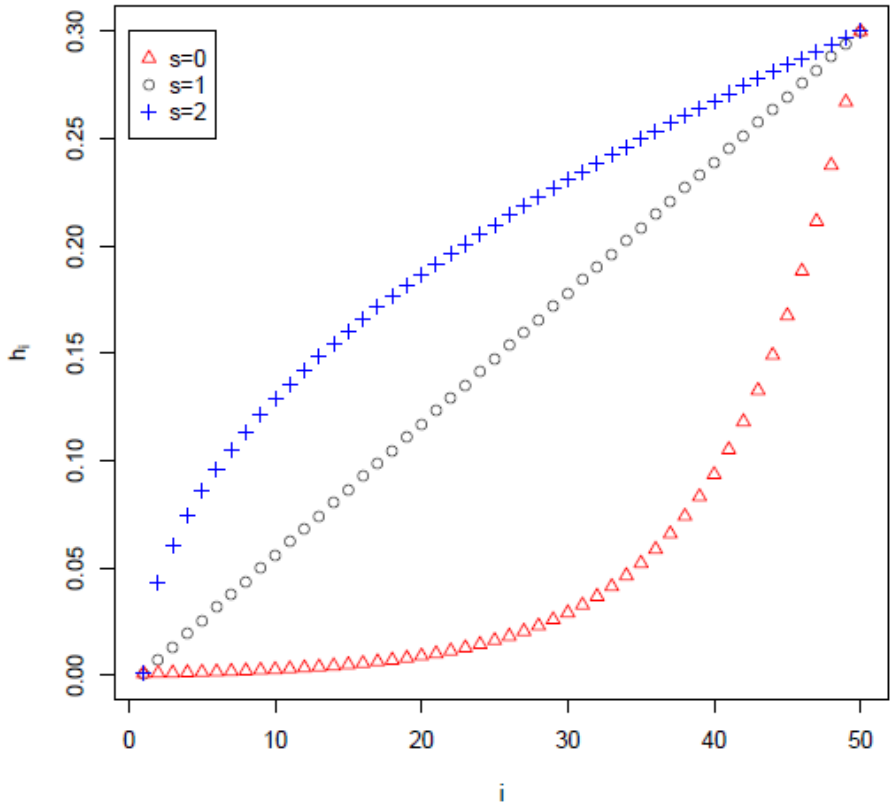
In this case, $\log h_i$ is obtained as a linear interpolation between $\log h_1$ and $\log h_m$. Thus, the parameter s can take any real value. Then, for $i = 1, 2, \ldots, m$,

$$h_i = \begin{cases} \left(\frac{m-i}{m-1} h_1^s + \frac{i-1}{m-1} h_m^s \right)^{1/s} & s \neq 0, \\ h_1^{\frac{m-i}{m-1}} h_m^{\frac{i-1}{m-1}} & s = 0. \end{cases} \tag{12}$$

This structure is reminiscent of the well-known Box–Cox transformation introduced by Box and Cox (1964).

Note that when $s = 1$, we have a linear pattern of h_i values with increasing i. When $s = 0$ we have an exponential pattern of h_i values. Furthermore, when $s < 1$, the pattern will be convex, and when $s > 1$, the pattern will be concave. Figure 2 shows the patterns of h_i values for $s = 0$, 1, and 2. For human aging, we would expect that $s = 0$ is the most appropriate of the three values. The value that one actually uses may be estimated from lifetime data.

Figure 2: Values of h_i determined using $h_1 = 0.001$, $h_m = 0.3$, and $s = 0$, 1, and 2.



In summary, we propose a Coxian distribution with five parameters to be specified: m, ψ, h_1, h_m, and s. We may be able to choose m and ψ based on prior knowledge, or a combination of lifetime data and prior knowledge. Having established the values of m and ψ, λ is determined from (11), and h_1, h_m, and s are easily estimated from lifetime data.

Ideally, our aging model would be calibrated using aging-related data, rather than just lifetime data. In other words, if observations of one or more key aging variables are taken at one or more times while each individual is alive in addition to ages at death, we could estimate the model parameters with less uncertainty. Govorun et al. (2018) describe a method of doing this. However, it is beyond the scope of this paper.

Since the states in our model are not connected to any aging variables, these states are unobservable. Therefore, it is not possible to identify the state of a specific individual and use this information in calculating probabilities related to his/her future lifetime. The usefulness of our model is more in capturing the aging mechanism and understanding the variability that this mechanism creates.

Traditional life table mortality rates can be calculated from our model using

$$q_x = \frac{S(x)-S(x+1)}{S(x)},$$

where values of the survival function $S(\cdot)$ are obtained from (5). These mortality rates increase with age. However, it is important to note that none of our model transition rates

increase with the age of an individual. What changes over time is the state of the individual. At older ages, probabilities associated with later aging states are higher. And since the state-specific death rates are higher in those later states, we have mortality rates that increase with age.

3. Model estimation and application

3.1. Parameter estimation

The parameters in our model can be estimated using maximum likelihood estimation, which we now describe.

Suppose that θ is the parameter vector we wish to estimate. If we have chosen values of m and ψ, then $\theta = (h_1, h_m, s)$. Further, suppose we wish to estimate this parameter vector using data on lifetimes only. Specifically, for each observed individual $j = 1, \ldots, n$ we observe (ℓ_j, t_j, δ_j), where ℓ_j is the age at which individual j entered observation, t_j is the age at which individual j left observation, and $\delta_j = 1$ if individual j died at time t_j and $\delta_j = 0$ otherwise.

The likelihood function is given by

$$L(\theta) = \prod_{j=1}^{n} \frac{f(t_j; \theta)^{\delta_j} S(t_j; \theta)^{1-\delta_j}}{S(\ell_j; \theta)}, \tag{13}$$

where $S(t; \theta)$ is our model survival function, and $f(t; \theta)$ is our model probability density function. Recall that we have

$$S(t; \theta) = \alpha^T \exp(\Lambda t)\, e,$$

and

$$f(t; \theta) = \alpha^T \exp(\Lambda t)\, q.$$

Both Λ and q are functions of θ. Therefore, for a given θ, we can calculate the value of the likelihood or the log-likelihood, $l(\theta) = \log L(\theta)$. The log-likelihood can then be maximized numerically with respect to θ to obtain the maximum likelihood estimates.

3.2. Application using data from a retirement community

To illustrate how our aging model can be calibrated using lifetime data, we consider data from Channing House, a retirement community in Palo Alto, California. The data set includes the entry age and age at death (or study end) for 462 people (97 males and 365 females) who resided in the facility between January 1964 and July 1975. These residents were covered by a health care program, which provided easy access to care at no cost. This may have resulted in lower than average mortality. We fit our aging model to the female data only. Thus, we consider data on a relatively small group of homogeneous individuals – all females living in the same community with the same access to health care and, very likely, similar lifestyles. This is ideal, as an assumption of homogeneity underlies our model.

Only 362 of the 365 female records were kept, because three records had equal entry and exit ages. Of the 362 females, 130 died while in the community, and the other 232 survived until the

end of the observation period. The youngest entry age was just over 61, and the oldest exit age was just under 101. So the data pertain to aging over this age range.

Since our model assumes all individuals start in state 1, but we expect some variability in state by age 61, we assume for this example that the aging process starts at age 50. This allows us to achieve some variability in the state distribution by age 61. Also, we assume, somewhat arbitrarily, that age 105 is the end of the life span; all members of the community were dead by this age. Therefore, $\psi = 105 - 50 = 55$. Also, based on experience, we believe that $m = 100$ aging states is appropriate for modelling aging above age 50. This results in a reasonable amount of variability in the physiological age at various calendar ages, as shown later in Figure 6. We then have that $\lambda = m/\psi = 100/55 = 1.8182$. Using maximum likelihood estimation as described above, we obtain $\hat{h}_1 = 0.00175471$, $\hat{h}_m = 1.27518$, and $\hat{s} = -0.0734710$. The parameter estimators are highly correlated and the variances are large. However, the estimates seem reasonable and lead to a good fit to the data.

The estimates of h_i that result from $\hat{h}_1$, $\hat{h}_m$, and $\hat{s}$ are shown in Figure 3. The pattern of values looks reasonable given that they represent instantaneous mortality rates that apply between ages 50 and 105.

Figure 3: Estimates of h_i obtained by calibrating the PTAM using the Channing House female data.

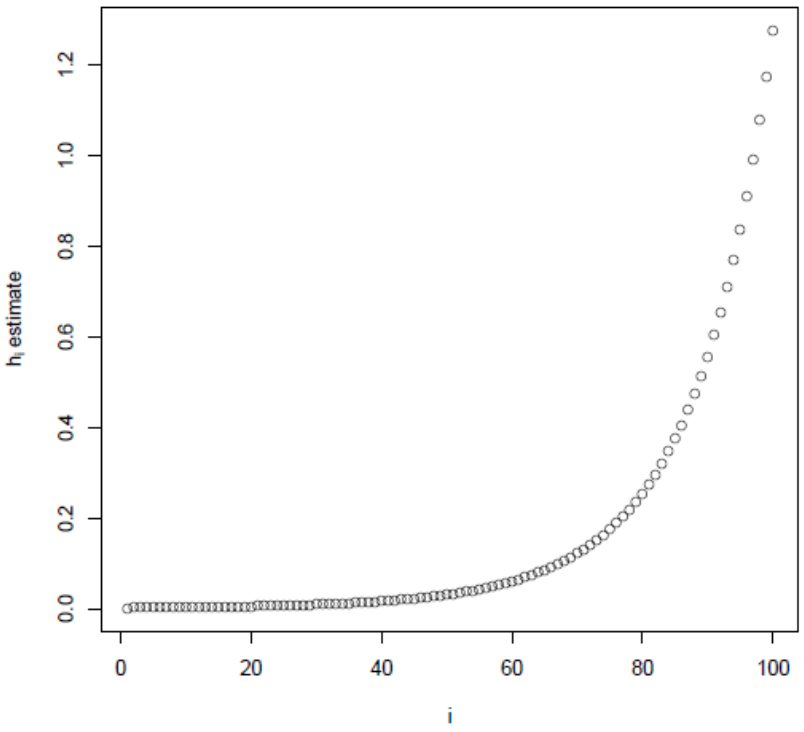
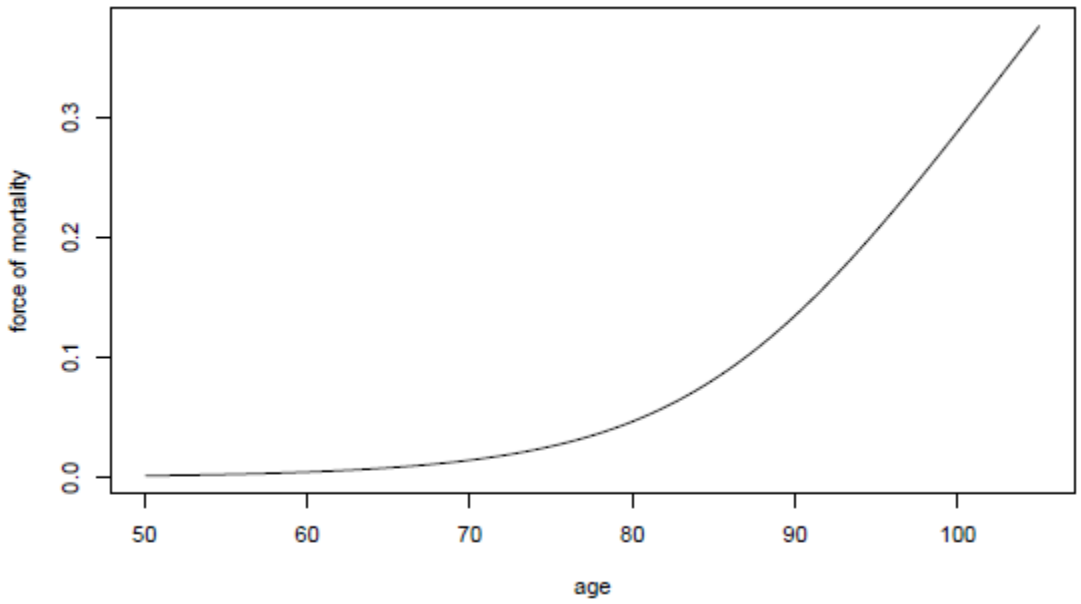
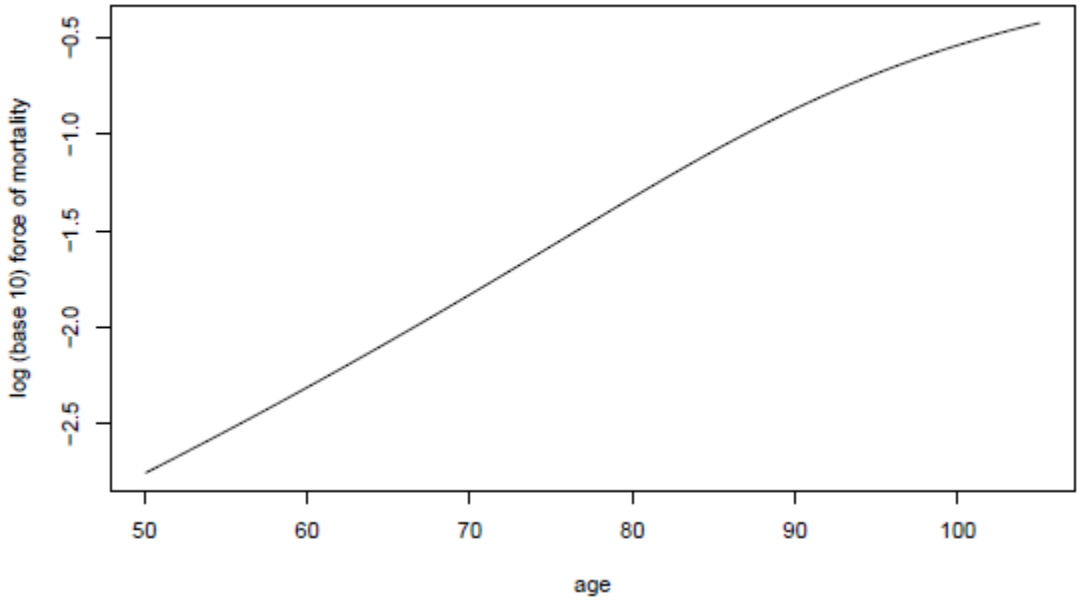


Figure 4 shows the force of mortality and the log (base 10) of the force of mortality based on the fitted model. Once again the behaviour is quite reasonable.

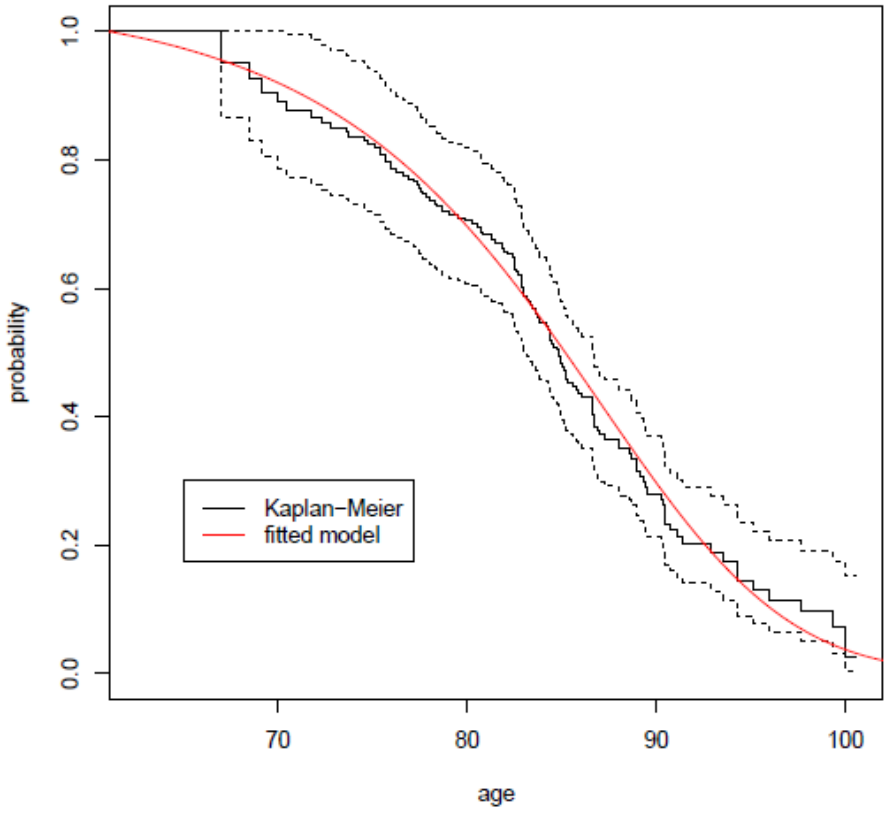
Figure 4: Force of mortality and log (base 10) force of mortality based on the PTAM calibrated using the Channing House female data.





In Figure 5 we illustrate the goodness of fit of our model to the Channing House data by plotting the fitted survival function along with the Kaplan–Meier nonparametric survival function estimates. We observe that our model fits quite well – and our fitted model estimates stay within the 95 per cent confidence limits based on the Kaplan–Meier estimator.

Figure 5: Survival function based on the PTAM calibrated using the Channing House female data, along with the Kaplan–Meier estimates of the survival function and corresponding 95 per cent confidence limits (dashed).



With our fitted model, we can perform a variety of analyses of the aging process. For example, it is interesting to examine the distribution of the state of an individual at different ages, given that the individual is alive at these ages. The probabilities of physiological age associated with this distribution, that is $\Pr(Y_t = i | Y_t \in E)$, are given by the vector

$$\mathbf{p}(t)/\mathbf{p}(t)e,$$

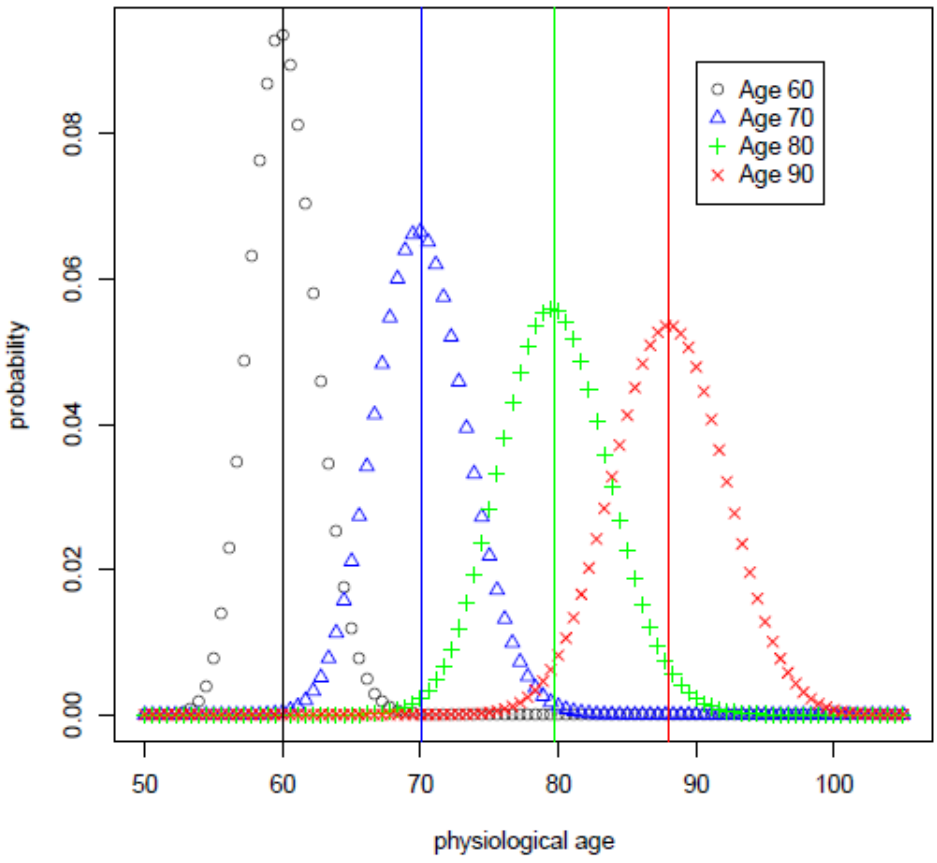
where $\mathbf{p}(t)$ can be calculated using equation (6). Since these probabilities correspond to age $50 + t$ in our example, it is intuitively appealing to transform the state Y_t to a comparable value that we can interpret as the individual's physiological age. Let

$$Physiological\ age\ at\ calendar\ age\ t\ = 50 + \frac{Y_t - 1}{m - 1}\,\psi.$$

Then an individual in state 1 has physiological age 50, and an individual in state 100 has physiological age $50 + \frac{100-1}{99}(55) = 105$. We can now determine the distribution of physiological age at any calendar age.

Figure 6 shows the distribution of physiological age at ages 60, 70, 80, and 90 for the PTAM calibrated using the Channing House female data. We observe that the physiological age distribution has less variability at younger ages. This is mainly due to the fact that we started the process at age 50, so that the all individuals have physiological age 50 at age 50. At older ages, the variability stabilizes. We also observe that the mean physiological age, indicated by the vertical lines, is very close to 60 at age 60. However, at older ages, the mean physiological age becomes noticeably smaller than the age. This is due to mortality selection – individuals with a higher physiological age have a higher mortality rate. Therefore, survivors tend to be those with a lower physiological age.

Figure 6: Physiological age distribution at ages 60, 70, 80, and 90 based on the PTAM calibrated using the Channing House female data. Vertical lines indicate the means of the distributions.

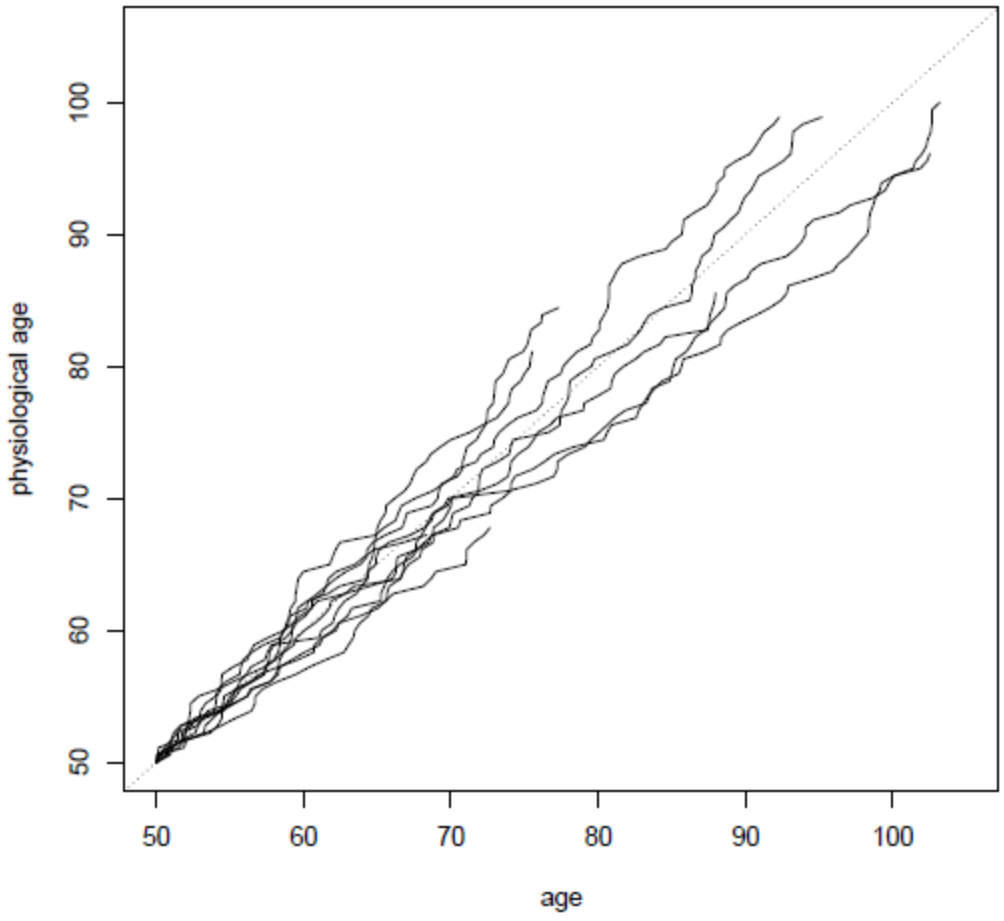


We can gain further insight into our aging model by observing how the paths of the process behave. We can do this by simulating several paths and plotting them. To simulate a path, we assume that the individual is in state 1 (physiological age 50) at age 50. The individual will stay in state 1 for a length of time that is exponentially distributed with rate $\lambda + h_1$. We can

generate this exponential random variable. At the end of the stay in state 1, the individual will die with probability $h_1/(\lambda + h_1)$ and move to state 2 with probability $\lambda/(\lambda + h_1)$. We can generate a uniform (0,1) random variable to determine whether or not the individual dies at this time. If not, we generate the time spent in state 2 (exponential with rate $\lambda + h_2$) and continue until the individual dies.

We simulated 10 paths of the process and plotted them in Figure 7. Once again the state of the process has been transformed to physiological age as described above. The figure illustrates once again the variability in physiological age at different ages, but also shows the extent to which individual paths depart from uniform aging over time. That is, we observe the "wigglyness" of individual paths.

Figure 7: Ten simulated paths from age 50 until death based on the PTAM calibrated using the Channing House female data.



4. Further analysis of our proposed PTAM

Our basic modelling principle is that the (length of) life is the result of two important forces: the force of aging (i.e., the intrinsic biological process) and the force of dying (i.e., the external forces from environmental stress including accidents, accessibility of nutrition and medical care, etc.). In addition, we want to model the two forces using a *stochastic* approach; specifically, we adopt the Markovian model framework of the Coxian distribution. That is, we define a finite number of states representing different physiological capacity levels and transition intensities determining the rate of progression through the states. These states (with labelling suitably transformed) are interpreted as an individual's physiological age, since they are used to classify an individual's physiological capacity at the moment. As a result, we have naturally incorporated the concept of heterogeneity into the lifetime dynamic process by introducing the so-called "*physiological age*".

We mentioned in Section 1 that, since the middle of the 20th century, many researchers have begun to collect longitudinal data on various physiological variables. The most striking finding may be the fact that the decline of these physiological functions as a result of the underlying aging process follows a slow, uniform, and roughly linear pattern with age. This suggests that, according to our assumption of a uniform rate of increase in physiological age, aging effect can be formulated as an increasing function of physiological age.

Notwithstanding the appeal of our approach, other strategies can be used to construct a Coxian distribution. We discuss one of these below and illustrate that our model is nearly equivalent in terms of the resulting lifetime distribution.

4.1. The Le Bras dual linear Markovian model

We now describe a model developed by Szilard (1959) and Le Bras (1976). Szilard proposes an aging process theory. It assumes that chromosomes mutate with constant rates in cells. If a cell accumulates too many mutations, it will cease functioning. Once a certain percentage of cells stop functioning in a human body, the body will die. Furthermore, Le Bras added an assumption that the inherited chromosomal mutation in the human body follows a Markov process. For a newborn, each cell mutates with rate λ_0 initially. Then new mutations occur with additional rate $i\lambda$, proportional to the total number of mutations i that have happened in this cell, and each cell dies with rate $i\mu$, proportional to i as well. Hence, let state i of a Markov process represent the state in which cells have accumulated i mutations in total. Then the transition rate from state i to $i+1$ is $\lambda_0 + i\lambda$ and the transition rate from state i to the absorbing state (i.e., to death) is $i\mu$.

Figure 8: Diagram for the Le Bras dual linear model.

The Le Bras model has been discussed by Yashin et al. (1994) in the following form: $\lambda_i = \lambda_0 + i\lambda$ and $\mu_i = \mu_0 + i\mu$, where μ_0 is the initial exiting rate. See Figure 8 for the diagram of the model. Here, the Le Bras model has an infinite number of states. Following the steps in Section 2, one can derive the transition probability

$$P_i(t) = \frac{e^{-(\mu_0+\lambda_0)t}}{i!}\left(\frac{\lambda(1-e^{-(\lambda+\mu)t})}{\lambda+\mu}\right)^i \prod_{k=1}^{i}\left(\frac{\lambda_0}{\lambda}+k-1\right). \quad (14)$$

Recognizing that $P_i(t)$ follows a binomial series pattern, it can be verified that

$$S(t) = \sum_{i=0}^{\infty} P_i(t) = e^{-(\lambda_0+\mu_0)t}\left(\frac{\lambda+\mu}{\mu+\lambda e^{-(\lambda+\mu)t}}\right)^{\frac{\lambda_0}{\lambda}}. \quad (15)$$

The Le Bras model is one of the first few attempts that have successfully incorporated physical/biological assumptions in a mathematical framework. Hence, it is important to review their approach, as this can help us to gain insight on how our proposed model can be useful.

More remarks about the Markovian approach and the Le Bras model:

- Under the further assumption that $\mu \ll \lambda$, the hazard function of the Le Bras model can be approximated by

  $$\mu(t) = \left(\mu_0 - \frac{\mu\lambda_0}{\lambda}\right) + \frac{\mu\lambda_0}{\lambda}\, e^{(\lambda+\mu)t},$$

  which is equivalent to the three-parameter Gompertz–Makeham mortality model, $\mu(t) = a + be^{ct}$, with

  $$a = \mu_0 - \frac{\mu\lambda_0}{\lambda}, \quad b = \frac{\mu\lambda_0}{\lambda}, \quad c = \lambda + \mu\,.$$

  See Yashin et al. (1994) for further details. The important contribution of the Le Bras model is the dual linear structure in describing the aging process (the linear pattern for λ_i) and the deteriorating effect (the linear pattern for μ_i) that could result in a Gompertz form of exponentially increased mortality pattern.

- In Yashin et al. (1994), the authors discuss the Le Bras model. It was found that, starting from a fixed frailty assumption, it is possible to derive the same mortality model. Hence, it was argued that, in the statistical analysis of data, results and conclusions depend not only on the data but also on basic assumptions about the mechanism that generated the data. In other words, in reality the use of lifetime data alone is not sufficient to distinguish between different mechanisms generating the observed mortality patterns. More sophisticated data need to be used to validate the assumption.

- As we discussed earlier, people die for many reasons, and the majority of causes are actually external. Therefore, it is hard to accept a model (e.g., the Le Bras model) that explains aging as the only cause of death. In many developed countries, another modern phenomenon of life is that advanced medical care now allows ordinary people to live a lot longer in a fragile physical status than before. We would like our model to be able to address the difference between the intrinsic force of aging and the external force of dying.

For all of these reasons, we prefer our model structure to that of Le Bras model.

4.2. A simulation study – from Le Bras model to the PTAM

We are interested in examining if our PTAM can capture the probabilistic features of the Le Bras model.

To explore this, we simulated 5,000 lifetime observations from the Le Bras model with the parameters given in Table 1. We then fit our model to the simulated data. Before estimating the other parameters using the MLE method, we set the life span parameter ψ as

$$\psi = \widehat{TVaR}_{0.999}(T), \tag{16}$$

where $\widehat{TVaR}_{1-\alpha}(T)$ is an empirical estimate of $TVaR_{1-\alpha}(T)$, obtained from the simulated data. Since our simulated sample size is 5,000, $\widehat{TVaR}_{0.999}(T)$ is the average of the five largest observations. We obtain $\psi = \widehat{TVaR}_{0.999}(T) = 112.55$. We estimate the remaining parameters using the MLE method. Figure 9 shows a histogram of the simulated data, the probability density function (pdf) of the Le Bras model and the pdf of the fitted PTAM. The MLE of m is determined by fitting the model for different but fixed m using the MLE method, and comparing their respective Negative Log-Likelihood (NLL) values. As one can see in Table 2, when $m = 225$, we obtain the lowest NLL value.

Table 1: Parameters value

The Le Bras model	λ_0	λ	μ_0	μ
	0.6	0.07	0.001	0.4×10^{-4}
Our fitted PTAM	h_1	h_m	λ	s
with $m = 225$	0.0008	1.65349	1.99908	-0.11118

Figure 9: Histogram of 5,000 lifetimes simulated from the Le Bras model. The fitted model with m = 225 is plotted along with the true model. The dotted vertical line indicates the location of $\psi = 112.55$.

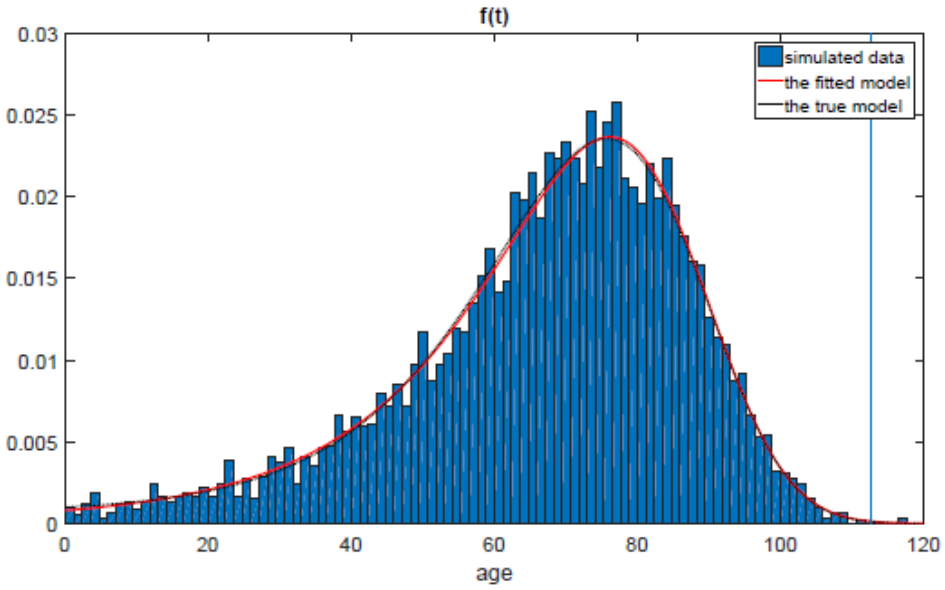


Figure 9 shows that if the observed mortality rates are truly generated from the Le Bras model, our model can provide a nearly equivalent representation. We are able to achieve the same goodness of fit as the Le Bras model by allowing the exiting rate h_i to increase faster than linearly. The fact that s can be flexible to take any value in order to accommodate the data (in this example $s = -0.11118$) is a specific feature of our model.

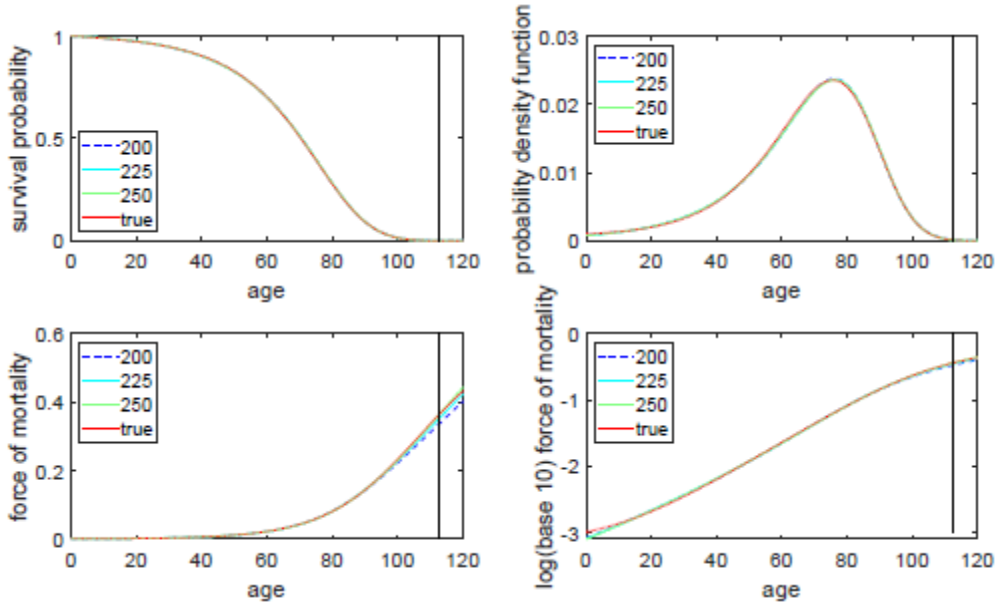
Hence, we have found an alternative to the Le Bras dual linear model with constant transition rate λ, though our model has a slightly different interpretation of the underlying aging mechanism. While the original Le Bras model contains an infinite number of the states, our PTAM re-labels them into m states. The aging process is still modelled as marching forward from one state to the next, and the impact of aging is still described as increased frailty to hazard with a higher indexed state. The only difference is the way of labelling the states. Transforming from the Le Bras model to our PTAM, the earlier states may need to be split into more states, while the later states may need to be grouped, but not in a linear fashion. The overall effect is that, in our model framework, the transitions from one to another are required to occur evenly due to the use of the constant transition rate λ. Correspondingly, the pattern for h_i changes from a linear increase to a pattern that has to climb slightly faster than exponentially, as indicated by the parameter s taking a small negative value.

In Table 2, we show additional results from fitting our model to the simulated data. As required, the estimate of the parameter λ increases with m. We have also plotted the survival functions, density functions, and hazard functions of these fitted models with m ranging from 200 to 250 in Figure 10; the dotted vertical line indicates the location of $\psi = 112.55$. For ages up to this value, the distributions are very close to each other. However, for the hazard functions there are small but noticeable differences beginning near age 100.

Table 2: Estimation results using different m based on 5,000 lifetimes simulated from the Le Bras limit distribution. The first column gives the NLL. The last column is the limit of the resulting hazard function $h(t)$ as $t \to \infty$.

NLL	h_1	h_m	λ	s	m	$\min(\lambda + h_1, h_m)$
21631.884	0.00081	2.00325	1.77637	-0.12373	200	1.77718
21631.826	0.0008	1.83922	1.86518	-0.11829	210	1.83922
21631.806	0.0008	1.70795	1.95400	-0.11336	220	1.70795
21631.713	0.0008	1.65349	1.99908	-0.11118	225	1.65349
21631.813	0.00079	1.60064	2.04282	-0.10886	230	1.60064
21631.843	0.00078	1.51076	2.13164	-0.10469	240	1.51076
21631.889	0.00078	1.43544	2.22046	-0.10089	250	1.43544

Figure 10: Fitted survival function - *S(t)*, probability density function - *f(t)*, hazard function - *h(t),* and log (base 10) hazard function. Each graph includes four curves corresponding to the fitted model with *m* = 200, 225, and 250, as well as the true model. The dotted vertical line indicates the location of $\psi = 112.55$.



Looking more carefully at the hazard functions in the bottom left panel of Figure 10, we find that the fit does not change symmetrically as m moves away from its optimal value. The model with $m = 250$ fits much better than the one with $m = 200$. Also, the model with $m = 250$ gives the highest hazard rate at the right end of the life span. For phase-type distributions, it is a well-known property that

$$\lim_{t\to\infty} h(t) = \min_{i=1,\cdots,m} \{d_1, d_2, \cdots, d_m\},$$

where the d_i's are the eigenvalues of the transition intensity matrix $\mathbf{\Lambda}$. In our PTAM, we have

$$\lim_{t\to\infty} h(t) = \min\{\lambda + h_1, h_m\},$$

since the eigenvalues $\lambda + h_i$ increase except for the last element of the matrix $\mathbf{\Lambda}$. The limit of $h(t)$ for each fitted model is provided in the last column of Table 2, and the model with $m = 250$ has the lowest value of this limit.

In Figure 11, we show the logarithm of the hazard function of the fitted models from age 80 to 500. While the graph shows interesting differences in tail behaviour, these differences occur well beyond the age range that is relevant from a practical perspective.

Figure 11: Fitted log (base 10) hazard function extended to age 500. Each graph includes four curves corresponding to the fitted model with *m* = 200, 225, and 250, as well as the true model. The dotted vertical line indicates the location of $\psi = 112.55$.

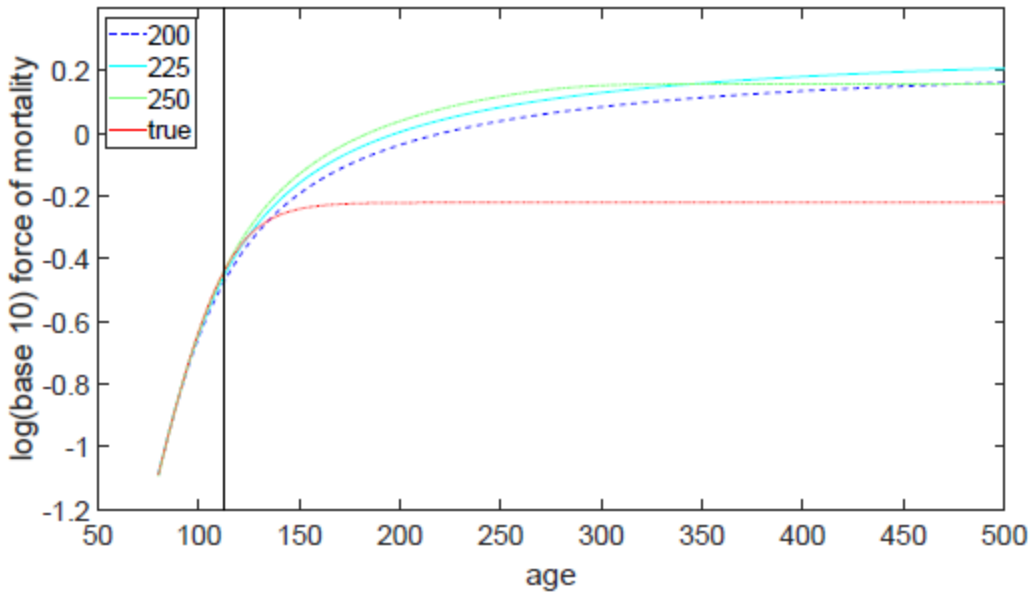
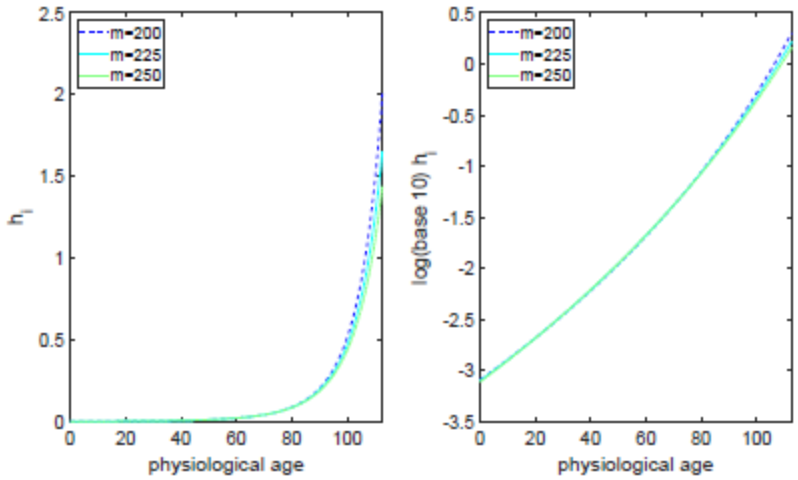


Figure 12: The left graph shows the exit rate h_i with m = 200, 225 and 250. The right graph shows the log-exit rate with m = 200, 225, and 250. Both are plotted against the physiological age $\frac{i-1}{m-1}\psi$.



In Figure 12, we plot h_i versus physiological age $\frac{i-1}{m-1}\psi$. The pattern of h_i on $[0,\psi]$ turns out to be quite stable with different m. This is another good feature of the proposed model, indicating that the underlying mechanism does not vary too much with m.

In our analysis so far, we have demonstrated that some key aspects of our aging model are quite stable with changing m. Our estimated value of m was determined by maximizing the likelihood. However, the likelihood is quite flat for values of m near the maximum. So there are only small differences in the lifetime distribution that result from different m. This leads to the question: how is the model affected by the value of m and how should m be determined?

While the lifetime distribution is relatively insensitive to m, the progression of physiological age with advancing calendar age is affected by m. In particular, the variability in physiological age is significantly affected by m. In fact, as $m \to \infty$, this variability disappears, and physiological age converges to calendar age. In particular, when m increases, the expected physiological age of an alive individual at calendar age $t \in (0,\psi)$ approaches $\frac{t}{\psi}m$ and the variability decreases. This phenomenon is illustrated in Figure 13, where 10 simulated sample paths for each of four different values of m – 25, 100, 225, and 1,000 – are plotted. In fact, we hypothesize that with $m \to \infty$, $\mathbb{E}(Y_t|Y_t \in E) = \frac{t}{\psi}m$ and $Var(Y_t|Y_t \in E) \to 0$.

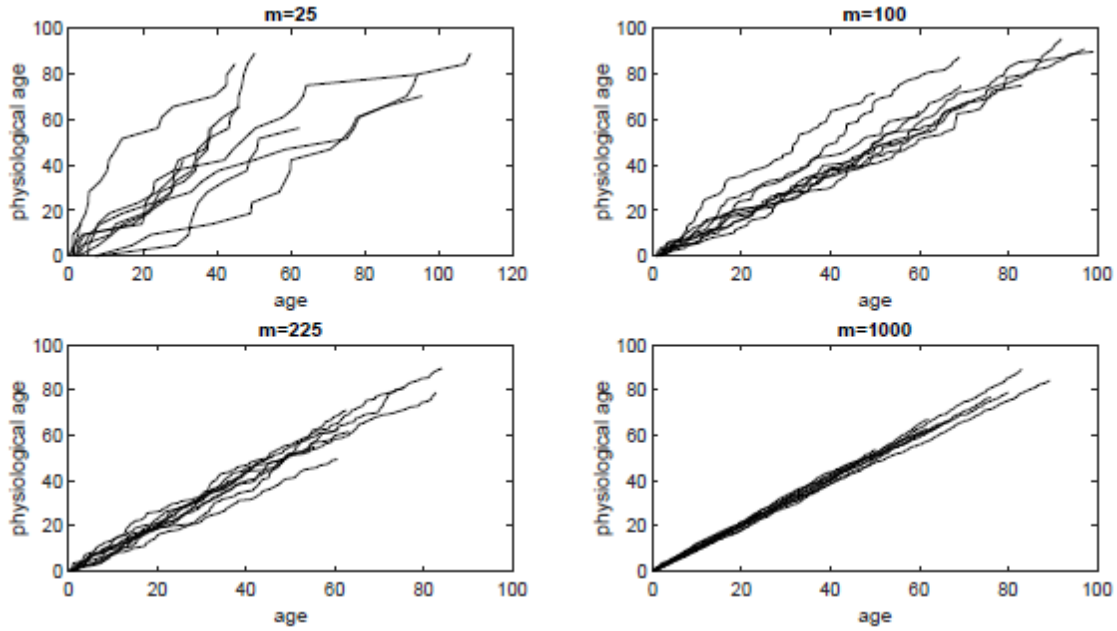
Furthermore, the resulting hazard function takes on the behaviour of our h_i values. That is, as m gets large (while keeping h_1, h_m, and s unchanged), we observe that the hazard function, which is the expected exit rate, given by $h^m(t) = \mathbb{E}(h_{Y_t}|Y_t \in E)$, approaches

$$h(t) = \begin{cases} \left(\left(1 - \frac{t}{\psi}\right) h_1^s + \frac{t}{\psi} h_m^s \right)^{1/s} & s \neq 0, \\ h_1^{1-\frac{t}{\psi}} \, h_m^{\frac{t}{\psi}} & s = 0. \end{cases} \tag{17}$$

In the same time, the variability of h_{Y_t} becomes smaller and smaller when m increases. This is an interesting property of the PTAM in which the role of parameter m is closely related to the degree of the uncertainty in determining aging and/or interpreting aging effect.

Figure 13 shows 10 simulated sample paths for each of four different values of m: 25, 100, 225, and 1,000. We observe that, as m increases, there is less variability in physiological age. And physiological age is converging to calendar age. This suggests that m should not be too large, or the model will not appropriately reflect the variability in physiological age. Obviously, m cannot be too small either, or there will too much variability in physiological age.

Figure 13: Ten simulated sample paths of the fitted PTAM for each of four different values of m, holding the parameters h_1, h_m, and s fixed.



This brings us back to the question of how to choose m. It can be chosen to reflect one's opinion about the variability in physiological age. However, if one has data that include information about one or more health variables related to physiological age in addition to the lifetimes of individuals, then these can be used to estimate m along with the other parameters. This is the preferred approach if suitable data are available.

5. Summary and conclusions

We have presented a mathematical model for aging. The model has a very natural interpretation, as individuals are assumed to age by progressing through a number of states, with a random time spent in each state. The model is made even more intuitive by transforming the state label to a scale that can be interpreted as physiological age. As one would expect, the mortality rate increases as one moves from one state to the next. However, the force of aging and the force of mortality are separate components of the model.

We assume that the rate of aging is constant over time, so that individuals progress uniformly through the states. The pattern of state-specific mortality rates is determined using a three-parameter structure that allows considerable flexibility and includes linear and exponential patterns as special cases. Once the number of model states is specified, the rate of aging is determined so that individuals will reach the final state only very close to the end of the life span, which must be specified.

Though the functional form chosen for the state-specific mortality rates is appealing, its flexibility is limited. Specifically, the form can only achieve a monotone pattern of rates with constant, increasing, or decreasing slope. This simple form cannot, therefore, capture a bathtub shape or an accident-related hump in the force of mortality. However, we believe it may be possible to combine functions of the form we propose in order to achieve greater flexibility. This is an avenue for future research.

Our model can be calibrated with lifetime data using maximum likelihood estimation. However, in this case we have a lot of uncertainty about the estimate of the number of model states. This makes sense because the number of states determines the variability in physiological age; as a result, we need data that carry relevant and explicit information about the aging process and its variation within the population of interest. More specifically, we would need longitudinal data of observed health-related variables which are good indicators for aging. At the current stage, since we have only used mortality rates to calibrate the model, the fitted model can only provide a relative measure of physiological age. Conditional on the availability of aging-related data, the model can be further enhanced to describe the state-specific aging characteristics. So one of the areas for future research is the exploration of methods for estimating parameters when data involve health observations as well as mortality information. The approach of Govorun et al. (2018) may be a good starting point for this development.

Another area for future research involves investigating how to modify the model to allow for different populations or different birth cohorts. One option is to use different aging rates with mortality rates that are the same. Another is to allow the mortality rate parameters to be different.

Acknowledgement

The authors gratefully acknowledge financial support for this research through a grant from the Canadian Institute of Actuaries.

References

Aalen, O. O. (1995). Phase-type distributions in survival analysis. *Scandinavian Journal of Statistics*, 22(4):447–463.

Asmussen, S. (1989). Exponential families generated by phase-type distributions and other Markov lifetimes. *Scandinavian Journal of Statistics*, 16(4):319–334.

Asmussen, S., Nerman, O., and Olsson, M. (1996). Fitting phase-type distributions via the EM algorithm. *Scandinavian Journal of Statistics*, 23(4):419–441.

Belsky, D. W., Caspi, A., Houts, R., Cohen, H. J., Corcoran, D. L., Danese, A., Harrington, H., Israel, S., Levine, M. E., Schaefer, J. D., et al. (2015). Quantification of biological aging in young adults. *Proceedings of the National Academy of Sciences*, 12(30):E4104–E4110.

Box, G. E., and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, 26(2):211–252.

Crimmins, E. M., Johnston, M., Hayward, M., and Seeman, T. (2003). Age differences in allostatic load: an index of physiological dysregulation. *Experimental Gerontology*, 38(7):731–734.

Dublin, L. I., Lotka, A. J., and Spiegelman, M. (1949). *Length of Life*. Ronald, New York.

Govorun, M., Jones, B. L., Liu, X., and Stanford, D. A. (2018). Physiological age, health costs, and their interrelation. *North American Actuarial Journal*, 22(3):323–340.

Herskind, A. M., McGue, M., Holm, N. V., Sörensen, T. I., Harvald, B., and Vaupel, J. W. (1995). The heritability of human longevity: a population-based study of 2872 Danish twin pairs born 1870–1900. *Human Genetics*, 97(3):319–323.

Jones, H. B. (1956). A special consideration of the aging process, disease, and life expectancy. *Advances in Biological and Medical Physics*, 4:281–337.

Le Bras, H. (1976). Lois de mortalité et âge limite. *Population*, 31(3):655–692.

Lin, X. S., and Liu, X. (2007). Markov aging process and phase-type law of mortality. *North American Actuarial Journal*, 11(4):92–109.

Neuts, M. F. (1982). Explicit steady-state solutions to some elementary queueing models. *Operations Research*, 30(3):480–489.

Slud, E. V., and Suntornchost, J. (2014). Parametric survival densities from phase-type models. *Lifetime Data Analysis*, 20(3):459–480.

Su, S., and Sherris, M. (2012). Heterogeneity of Australian population mortality and implications for a viable life annuity market. *Insurance: Mathematics and Economics*, 51(2):322–332.

Szilard, L. (1959). On the nature of the aging process. *Proceedings of the National Academy of Sciences*, 45(1):30–45.

Yashin, A. I., Arbeev, K. G., Ukraintseva, S. V., Akushevich, I., and Kulminski, A. (2012). Patterns of aging-related changes on the way to 100. *North American Actuarial Journal*, 16(4):403–433.

Yashin, A. I., Vaupel, J. W., and Iachine, I. A. (1994). A duality in aging: the equivalence of mortality models based on radically different concepts. *Mechanisms of Ageing and Development*, 74(1–2):1–14.